

Durham Research Online

Deposited in DRO:

31 July 2017

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Harrison, G.W. and Lau, M. and Ross, D. and Swarthout, J.T. (2017) 'Small stakes risk aversion in the laboratory : a reconsideration.', *Economics letters.*, 160 . pp. 24-28.

Further information on publisher's website:

<https://doi.org/10.1016/j.econlet.2017.08.003>

Publisher's copyright statement:

© 2017 This manuscript version is made available under the CC-BY-NC-ND 4.0 license
<http://creativecommons.org/licenses/by-nc-nd/4.0/>

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Small Stakes Risk Aversion in the Laboratory: A Reconsideration

by

Glenn W. Harrison, Morten Lau, Don Ross and J. Todd Swarthout[†]

July 2017

Forthcoming, *Economics Letters*

Abstract

Evidence of risk aversion in laboratory settings over small stakes leads to *a priori* implausible levels of risk aversion over large stakes under certain assumptions. One core assumption in statements of this calibration puzzle is that small-stakes risk aversion is observed over all levels of wealth, or over a “sufficiently large” range of wealth. Although this assumption is viewed as self-evident from the vast experimental literature showing risk aversion over laboratory stakes, it actually requires that lab wealth be varied for a given subject as one evaluates the risk attitudes of the subject. We consider evidence from a simple design that tests this assumption, and find that the assumption is strikingly *rejected* for a large sample of subjects from a population of college students. We conclude that the implausible predictions that flow from these assumptions do not apply to one specialized population widely used to study economic behavior in laboratory experiments.

Keywords: risk aversion, calibration puzzle, laboratory experiments

JEL codes: D81, D03, C91

[†] Department of Risk Management & Insurance and Center for the Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Harrison); Copenhagen Business School, Denmark (Lau); School of Sociology and Philosophy, University College Cork, Ireland; School of Economics, University of Cape Town, South Africa; and Center for Economic Analysis of Risk, Robinson College of Business, Georgia State University, USA (Ross); Department of Economics and Experimental Economics Center, Andrew Young School of Policy Studies, Georgia State University, USA (Swarthout). Harrison is also affiliated with the School of Economics, University of Cape Town. E-mail contacts: gharrison@gsu.edu, mla.eco@cbs.dk, don.ross@uct.ac.za and swarthout@gsu.edu. We are grateful to the Danish Social Science Research Council (Project #12-130950) for financial support, and to Jim Cox and Vjollca Sadiraj for valuable comments and discussion.

Debate surrounding theories of decisions under risk and uncertainty has renewed interest in the arguments of the utility function over event outcomes. The local measure of risk aversion proposed by Arrow [1971] and Pratt [1964] for expected utility theory (EUT) is based on terminal wealth being the argument. However, there is nothing in the axiomatic foundation of EUT that requires one to use terminal wealth as the argument: Vickrey [1945] used income instead of terminal wealth; von Neumann and Morgenstern [1953; p. 15-31] were agnostic; and Luce and Raiffa [1957; ch.2] discussed alternatives such as scalar amounts of terminal wealth or income or, alternatively, vectors of commodities. Arrow [1964], Debreu [1959; ch.7] and Hirshleifer [1966] developed models in which the arguments of utility functions are vectors of contingent commodities.

The choice of arguments of the utility function can have important consequences for the inferences one can plausibly draw from empirical estimates of risk attitudes. Many economics experiments present participants with gambles over relatively small stakes and find that such gambles are frequently turned down in favor of less risky gambles with smaller expected values. If the argument of the utility function is terminal wealth, then some patterns of small stakes risk aversion have implausible implications for preferences over gambles where the stakes are no longer small. One example from Rabin [2000] is that the expected utility of terminal wealth model implies that an agent who turns down a 50/50 bet of losing \$100 or gaining \$110, at all initial wealth levels between \$100 and \$300,000, will also, at initial wealth of \$290,000, turn down a 50/50 bet with possible loss of \$2,000 even when the gain is as large as \$12 million. Although primarily used as an argument against EUT, it is now well-known that this logic applies to a much wider range of models that assume the argument of the utility function to be terminal wealth (Cox and Sadiraj [2006] and Safra and Segal [2008]). Hence the methodological implications run much deeper than whether EUT is a useful descriptive model of behavior.

Given the importance of understanding the arguments of the utility function, the absence of empirical tests is remarkable. We focus squarely on the specific and influential claim that evidence of small stakes risk aversion in the laboratory generate implausible risk aversion implications for any model in which terminal wealth is the argument of the utility function.¹ We refer to this claim as the HRC, following Hansson [1988] and Rabin [2000]. We build on a simple theoretical test and experimental design originally developed by Cox and Sadiraj [2008], and independently later by Wilcox [2013]; our design follows theirs. Although this design has wider implications, we focus on implications for calibration puzzles defined over terminal wealth models, which are the models that initiated the modern debate.

We present direct evidence that *the empirical premiss underlying the HRC claim is false for the typical subjects of laboratory experiments*: students in a first-world university.² These subjects exhibit risk aversion for small stakes lotteries for the initial terminal wealth that they bring to the lab, but as we increase the terminal wealth of the subjects they exhibit risk neutrality. We make no claim that this finding generalizes to other populations, fully expect that it could vary from population to population, and encourage tests with different populations.³

We review the theoretical statement of the usual calibration puzzle in section 1, using the simple example from Hansson [1988] since it is not widely known and illustrates the basic points. The generalization by Rabin [2000] can then be quickly stated. Our experimental design is presented in

¹ Neilson [2001] and Safra and Segal [2008] also provide concavity calibration claims for terminal wealth models.

² There have been comparable tests of the premisses of the calibration claims by Cox, Sadiraj, Vogt and Dasgupta [2013]. One of their experiments involved subjects in Calcutta, India; another involved a casino in Europe and some experimental procedures that are non-standard (in effect, the lab wealth was extremely risky wealth, and plausibly hypothetical wealth from the *ex ante* perspective of the subject). Wilcox [2013] independently derived similar tests in the laboratory.

³ We now incorporate these simple lottery choices in most batteries we use in the lab and the field for new populations.

section 2, and follows Cox and Sadiraj [2008]. We evaluate the resulting empirical evidence from 28 binary choices by 590 student subjects in section 3.

1. Theory

What if an individual always rejects a 50:50 lottery offering x and $x+\$3$ in favor of $x+\$1$ for certain? Without loss of essential generality, assume indifference, so that $u(x+1) = \frac{1}{2} u(x) + \frac{1}{2} u(x+3)$. One solution to this equation is the utility function $u(x) = 1-a^x$ for $a \approx 0.618$. It is useful in the sequel to note that this is a bounded function, since $u(x) \rightarrow 1$ as $x \rightarrow \infty$. Now consider increments in utility from x :

$$u(x+1) - u(x) \approx 0.382 a^x \quad (1)$$

$$u(x+\infty) - u(x) = 1 - (1-a^x) = a^x \quad (2)$$

If (1) and (2) are true, then we can construct “trick lotteries” that make this decision maker look silly.

For instance, the decision maker must prefer a certain gain of 1 to a $0 < p < 0.382$ chance of an arbitrarily large gain $X \gg x$. For instance, $p = \frac{1}{4}$, since $\frac{1}{4} a^x < 0.382 a^x$, and $X = \$1$ million.

More general conditions for this implausible prediction are now established. If the utility function is bounded on $(0, \infty)$ then that is a sufficient condition for implausible risk aversion in large stakes (e.g., Cox and Sadiraj [2008; Proposition 2, p.20]). Indeed, the only empirical example offered by Rabin [2000] uses a bounded CARA function. On the other hand, small-stakes risk aversion over a *large enough finite* interval is a sufficient condition for implausible risk aversion for large stakes, whether or not the utility function is bounded or unbounded.

The Hansson-Rabin calibration (HRC) puzzle may be stated in terms of four propositions:

- P = “the agent is a risk averse EUT maximizer”
- Q = “EUT implies full asset integration”
- R = “the agent turns down small-stakes gambles in favor of a certain amount with a slightly lower expected value, and does so over a large enough⁴ range of wealth levels W ”

⁴ The expression “large enough” is deliberately vague, since it depends on the degree of risk aversion exhibited under proposition P, and the lotteries in proposition S that *a priori* seem to generate silly behavior.

- $S \equiv$ “the agent turns down large-stakes gambles in favor of a certain amount with a significantly lower value, and looks silly.”

The calibration puzzle is the claim that if P, Q and R are true, then S follows. Since the behavior implied by proposition S is *a priori* implausible from a thought experiment, there must be some inconsistency among these propositions. Rabin [2000] draws the implication that P must then be false, and that one should employ models of decision-making under risk that relax proposition Q, such as Cumulative Prospect Theory. As a purely logical matter, of course, this is just one way to resolve this calibration puzzle.

2. Experimental Design

All of the evidence claimed to support the premiss that decision makers in experiments exhibit small stakes risk aversion for a large enough finite interval comes from designs in which subjects come to the lab with varying levels of wealth and are faced with small-stakes lotteries. This is actually indirect evidence, even if it might be suggestive, since we do not know that different decision-makers have varying levels of wealth, and there is nothing in EUT that would lead one to assume that they have the same utility function. What is needed is an experimental design that varies the wealth of a given decision-maker, who can be presumed to behave consistently with one utility function during the lab session.⁵

Cox and Sadiraj [2008; p.33] propose an elegant design to test this claim correctly. Give the agent a choice between a safe lottery of w for sure, and a risky lottery of a 50:50 chance of $w-x$ or $w+y$, where $w-x \geq 0$ and $y > x > 0$. The key idea is to vary w in the lab.⁶ Consider values of w that can be

⁵ Thought experiments along these lines were developed by Watt [2002] and Palacios-Huerta and Serrano [2006], and showed that the implied risk aversion underlying proposition R are *a priori* implausible. Despite minor errors in some of the calculations of the latter, noted at <http://www.econ.brown.edu/Faculty/serrano/disclaimers/2006EL91dis.html>, the message does not change.

⁶ Cox and Sadiraj [2008; p.32] explain why one cannot test proposition R by only asking one lottery choice question of this kind at only one level of lab wealth, as in Barberis, Huang and Thaler [2006; p. 1071] and Schechter [2007]. In response to Watt [2002], Rabin and Thaler [2002; p.230] make exactly this mistake in

denoted w, w, w, w, w, W, W , where smaller values of the letter w denote smaller values of lab wealth.

Consider values like these for lab wealth that may be plausibly much less than the W that the subject has in the field prior to the experiment. The experimenter does not need to know W for a given subject, but by varying “lab wealth” for that subject the experimenter has considered small-stakes lottery choices over a loss of x and a gain of y for “field + lab” wealth levels $W+w, W+w, W+w, W+w, W+w, W+w$ and $W+W$, respectively, for that subject. This step of the design presumes that we vary lab wealth for a given subject, since then we can presume that field wealth W is constant for the experimental session.⁷ This step of the experimental design also assumes proposition Q, that the agent perfectly asset integrates field wealth and lab wealth.

If the agent prefers the safe lottery over the risky lottery for all of the lab wealth values the experimenter’s budget can afford, then we have verification of proposition R, at least for the range of “field + lab” wealth proscribed by the experimenter’s budget. If we observe the agent choosing the safe lottery for small levels of lab wealth but the risky lottery for larger levels of lab wealth, then proposition R is rejected for that agent. Of course, we do not expect deterministic patterns of choice, so one ought to make some claim about the statistical significance of these choice patterns. This is why we have 28 choices for each subject. Another nice feature of this experimental design is that we need not structurally model the EUT decision process for the agent: we can rely on simple statistical models such as (panel) probit, conditioned on lab wealth.

misunderstanding the existing experimental literature: “We refer any reader who believes in risk neutrality to pick up virtually any experimental test of risk attitudes. Dozens of laboratory experiments show that people are averse to far more favorable bets for smaller stakes. The idea that people are not risk neutral in playing for modest stakes is uncontroversial; indeed, nobody to our knowledge interprets the existing evidence as arguing that expected-value maximization (risk neutrality) is a good fit.” The tests proposed here address the hypothesis of expected-value maximization for the relevant theoretical proposition R.

⁷ In other words, one cannot implement this aspect of the design on a between-subjects basis, or the theoretical premiss could be violated by heterogeneity.

Table 1 shows the 28 lottery pairs we used. The smallest lab wealth was \$15 and the largest lab wealth was \$135. Table 1 shows that 20 of the choices involve 50:50 risky lotteries, and 8 of the choices involve 90:10 risky lotteries to allow relatively high outcomes for the risky lottery. The order of lotteries in Table 1 is solely for pedagogic convenience: the order of presentation was randomized independently for each subject. One lottery is the safe lottery, and the other lottery is the risky lottery. The far right columns of Table 1 show the Expected Value (EV) of each choice, and the difference in favor of the risky lottery. Since the risky lottery provided an EV that always slightly exceeded the EV of the safe lottery, we know that a choice of the safe lottery implies risk aversion under EUT. Since the EV difference is strictly positive, choices of the risky lottery are consistent with risk neutrality or risk preference under EUT.

The first 14 lottery choices reflect lower levels of lab wealth between \$15 and \$30, and the last 14 lottery choices, shaded in Table 1, reflect higher levels of lab wealth between \$120 and \$135. There are repetitions of the $-x$ and $+y$ values at different levels of lab wealth: hence the design builds in controls for the same losses and gains relative to lab wealth, and only varies lab wealth. For instance, choice pairs 2 and 26 have $x = 8$ and $y = 10$, with lab wealth of \$15 and \$135, respectively. So apart from the change in lab wealth, these are the same risky lotteries, as required by proposition R. Similarly, choice pairs 9 and 18 have $x = 18$ and $y = 21$ with lab wealth of \$25 and \$125, respectively. There is variation in the values of x and y , and the difference between x and y : the smallest difference between x and y is \$1 (e.g., choice pairs 3, 6, 7) and the largest difference is \$85 (choice pairs 1, 12, 22 and 27).

3. Evidence

The battery in Table 1 was administered, in random order, to 590 undergraduates recruited from the Georgia State University population. This battery was administered as part of a larger experiment, in which subjects made a total of 100 binary choices, spanning gain, loss and mixed frames.

A typical display is shown in Figure 1.⁸ Indifference was not permitted as an option. Subjects also completed a demographics survey, and subsequently participated in a “lab casino” in which choices mimicked those found in video slot machines. Each subject received a \$5 show-up payment. One feature of this experiment is that participation times in a given session were staggered, so that several subjects joined the session at 30-minute intervals. All instructions were then administered in written form, and with a video presentation on each computer that the subject listened to privately with headphones. This procedure was used to avoid subjects feeling any peer pressure to finish early or late. At the end of the 100 lottery choices, one choice was selected at random with physical dice, and then the lottery chosen was played out with physical dice.⁹ Subjects were paid privately. The instructions are reproduced in an appendix (available on request).

Figure 2 displays the basic results, using a simple (random effects, panel) probit specification in which the binary choice is the dependent variable and the explanatory variable is lab wealth. For the probit specifications we use quadratic and cubic terms to capture possible non-linearity in the estimated latent index. Standard errors allow for clustering by subject. The solid line shows the average prediction, and the shaded area shows the 95% confidence interval around that prediction. The results are clear: subjects exhibit some risk aversion at low levels of lab wealth and exhibit risk neutrality (or risk preference) at higher levels of lab wealth. Although relying on interpolation, the switch point

⁸ These lottery choices were embedded in a larger battery that included some choices in which there were losses out of house endowments displayed for those choices. When there were no losses, the display did not mention any house endowment, as in Figure 1.

⁹ We opted for using the Random Lottery Incentive Method (RLIM), where one of the 100 choices was chosen at random for playing out and payment. We did so for two reasons, recognizing that this is not as innocent a procedure as some maintain (see Harrison and Swarthout [2014], Cox, Sadiraj and Schmidt [2015] and Brown and Healy [2016] for evidence and extensive literature discussion). The first reason was to ensure that we collected choices over a wide enough array of lotteries to be able to test the premiss using within-subject data. If we had opted for giving one choice to each subject, to avoid using the RLIM, this would have been infeasible and would have required an extreme assumption of preference homogeneity across subjects. The second reason was that the null hypothesis being tested is normally stated assuming EUT, and RLIM is valid under EUT.

appears to be around a lab wealth of \$60. By this evidence we infer that proposition R is false for this population and these lottery choices.¹⁰

Table 2 lists estimates of the probability of safe choices using several alternative statistical specifications. One specification is a standard probit model, and another specification is a random effects probit model; both use quadratic and cubic terms in lab wealth. A third specification is the semi-nonparametric (SNP) approach of Gallant and Nychka [1987]. Although less parametric than the probit specifications, the SNP estimates do not account for the panel nature of these data. Our preferred specification is the random effects, panel probit, shaded in Table 2. The same qualitative conclusion applies for all three specifications. The random effects probit implies more extreme predictions above and below $\frac{1}{2}$ than the other two specification, but in all cases the 95% confidence intervals lie above $\frac{1}{2}$ for small levels of lab wealth, and lie below $\frac{1}{2}$ for higher levels of lab wealth.

4. Conclusions

For a sample from a population that is similar to populations that are widely used to study behavior in experiments, the empirical basis for the calibration puzzle does not exist.

¹⁰ The qualitative results are even stronger if one drops the “skewed” risky lotteries, which are those numbered 1, 5, 11, 12, 17, 21, 22 and 27 in Table 1.

Table 1: Experimental Design

All currency values in U.S. Dollars

ID #	Lab Wealth	Risky Lottery				-x	y	Expected Value		
		Prob ^{low}	Payoff ^{low}	Prob ^{hi}	Payoff ^{hi}			Safe	Risky	Difference
1	15	0.9	5	0.1	110	-10	95	15	15.5	0.5
2	15	0.5	7	0.5	25	-8	10	15	16	1
3	15	0.5	7	0.5	24	-8	9	15	15.5	0.5
4	20	0.5	10	0.5	32	-10	12	20	21	1
5	20	0.9	15	0.1	70	-5	50	20	20.5	0.5
6	20	0.5	10	0.5	31	-10	11	20	20.5	0.5
7	20	0.5	2	0.5	39	-18	19	20	20.5	0.5
8	25	0.5	12	0.5	40	-13	15	25	26	1
9	25	0.5	7	0.5	46	-18	21	25	26.5	1.5
10	25	0.5	10	0.5	43	-15	18	25	26.5	1.5
11	30	0.9	25	0.1	90	-5	60	30	31.5	1.5
12	30	0.9	20	0.1	125	-10	95	30	30.5	0.5
13	30	0.5	18	0.5	45	-12	15	30	31.5	1.5
14	30	0.5	15	0.5	46	-15	16	30	30.5	0.5
15	120	0.5	108	0.5	135	-12	15	120	121.5	1.5
16	120	0.5	110	0.5	132	-10	12	120	121	1
17	120	0.9	115	0.1	180	-5	60	120	121.5	1.5
18	125	0.5	107	0.5	146	-18	21	125	126.5	1.5
19	125	0.5	110	0.5	143	-15	18	125	126.5	1.5
20	125	0.5	112	0.5	140	-13	15	125	126	1
21	130	0.9	125	0.1	180	-5	50	130	130.5	0.5
22	130	0.9	120	0.1	225	-10	95	130	130.5	0.5
23	130	0.5	112	0.5	149	-18	19	130	130.5	0.5
24	130	0.5	120	0.5	141	-10	11	130	130.5	0.5
25	130	0.5	115	0.5	146	-15	16	130	130.5	0.5
26	135	0.5	127	0.5	145	-8	10	135	136	1
27	135	0.9	125	0.1	230	-10	95	135	135.5	0.5
28	135	0.5	127	0.5	144	-8	9	135	135.5	0.5

Figure 1: Binary Choice Display

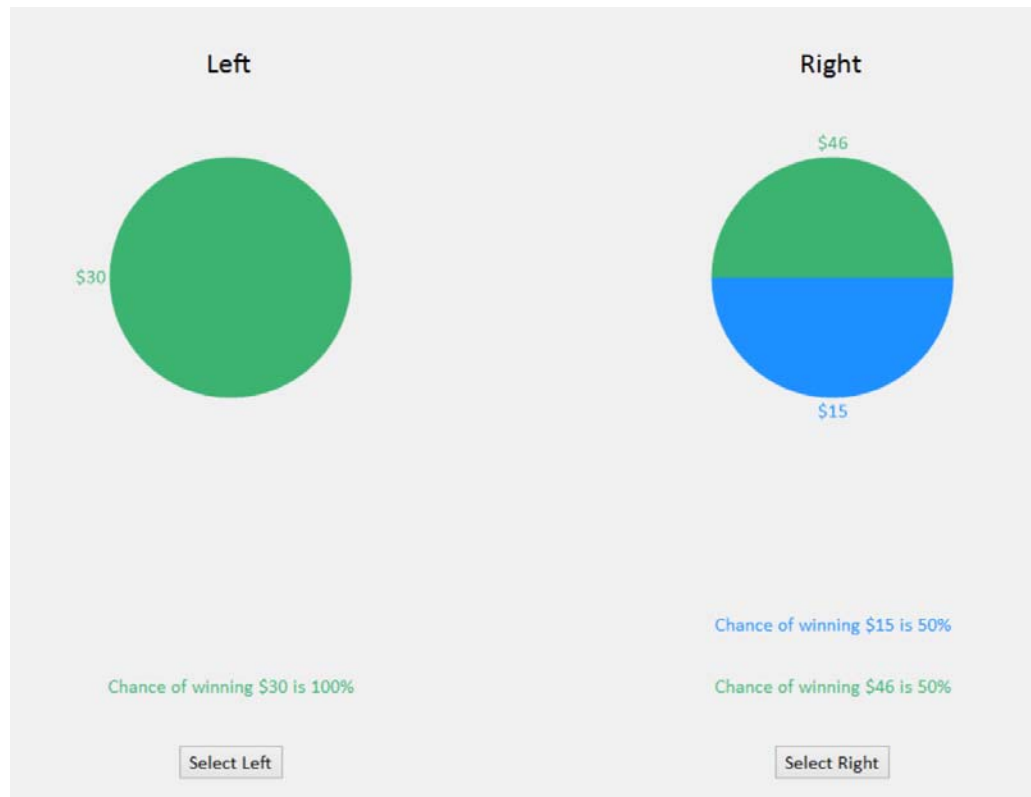


Figure 2: Predicted Probability of a
Risk Averse Choice
at Varying Levels of Lab Wealth

N=590 subjects each making 28 choices at each lab wealth
Predictions from a Random Effects Panel Probit model

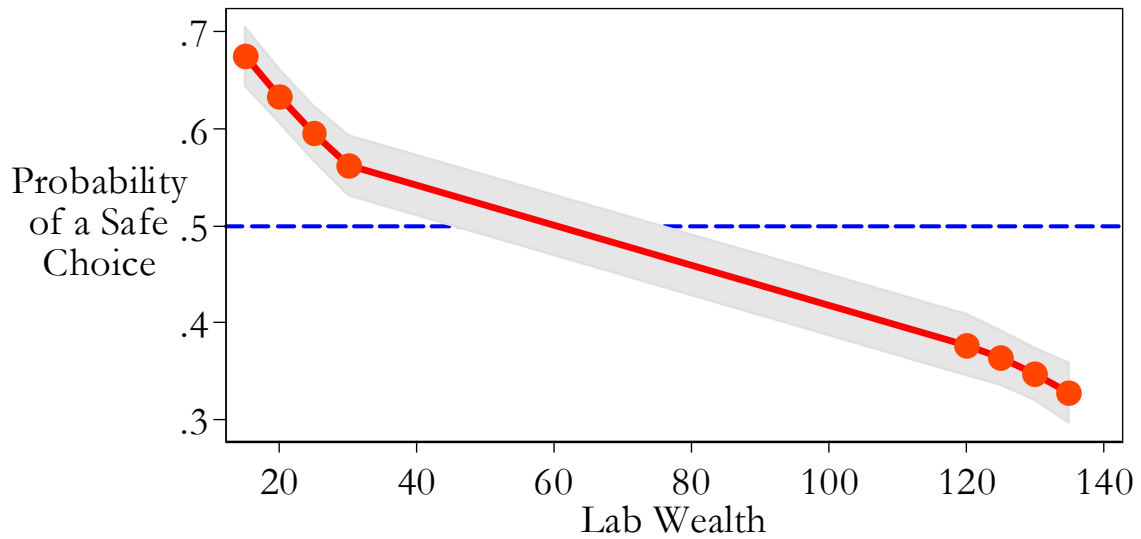


Table 2: Predicted Safe Choices from Alternative Statistical Models

Lab Wealth	Probit			Random Effects Probit			Semi-Non-Parametric (SNP)		
	Mean	95% CI		Mean	95% CI		Mean	95% CI	
\$15	0.64	0.62	0.67	0.68	0.65	0.71	0.64	0.62	0.66
\$20	0.61	0.59	0.63	0.63	0.61	0.66	0.61	0.60	0.62
\$25	0.58	0.55	0.60	0.60	0.57	0.62	0.58	0.57	0.59
\$30	0.55	0.53	0.57	0.56	0.53	0.59	0.55	0.53	0.56
\$120	0.40	0.37	0.43	0.38	0.35	0.41	0.40	0.38	0.42
\$125	0.39	0.36	0.41	0.36	0.34	0.39	0.39	0.38	0.40
\$130	0.38	0.35	0.40	0.35	0.32	0.37	0.38	0.36	0.39
\$135	0.36	0.33	0.39	0.33	0.30	0.36	0.36	0.34	0.38

References

- Arrow, Kenneth J., "The Role of Securities in the Optimal Allocation of Risk-Bearing," *Review of Economic Studies*, 31, 1964, 91-96.
- Arrow, Kenneth J., *Essays in the Theory of Risk-Bearing* (Chicago: Markham, 1971).
- Barberis, Nicholas; Huang, Ming, and Thaler, Richard H., "Individual Preferences, Monetary Gambles, and Stock Market Participation: A Case for Narrow Framing," *American Economic Review*, 96(4), September 2006, 1069-1090.
- Brown, Alexander L., and Healy, Paul J., "Separated Decisions," *Working Paper #16-02*, Department of Economics, Ohio State University, 2016.
- Cox, James C., and Sadiraj, Vjollca, "Small- and Large-Stakes Risk Aversion: Implications of Concavity Calibration for Decision Theory," *Games and Economic Behavior*, 56, 2006, 45-60.
- Cox, James C., and Sadiraj, Vjollca, "Risky Decisions in the Large and in the Small: Theory and Experiment," in J. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).
- Cox, James C.; Sadiraj, Vjollca, and Schmidt, Ulrich, "Paradoxes and Mechanisms for Choice under Risk," *Experimental Economics*, 18, 2015, 215-250.
- Cox, James C.; and Sadiraj, Vjollca; Vogt, Bodo, and Dasgupta, Utteeyo, "Is There a Plausible Theory for Decision under Risk? A Dual Calibration Critique," *Economic Theory*, 54(2), October 2013, 305-333.
- Debreu, Gerard, *Theory of Value* (New York: John Wiley and Sons, 1959).
- Gallant, A.R., and Nychka, Douglas W., "Semi-nonparametric maximum likelihood estimation," *Econometrica*, 55, 1987, 363-390.
- Hansson, Bengt, "Risk Aversion as a Problem of Conjoint Measurement," in P. Gardenfors and N-E. Sahlin (eds.), *Decisions, Probability, and Utility* (New York: Cambridge University Press, 1988).
- Harrison, Glenn, and Swarthout, J. Todd, "Experimental Payment Protocols and the Bipolar Behaviorist," *Theory and Decision*, 77, 2014, 423-438.
- Hirshleifer, Jack, "Investment Decision Under Uncertainty: Applications of the State-Preference Approach," *Quarterly Journal of Economics*, 80(2), 1966, 252-277.
- Luce, R. Duncan, and Raiffa, Howard, *Games and Decisions: Introduction and Critical Survey* (New York: Wiley, 1957).
- Neilson, William S., "Calibration Results for Rank-Dependent Expected Utility," *Economics Bulletin*, 4, 2001, 1-5.
- Palacios-Huerta, Ignacio and Serrano, Roberto, "Rejecting Small Gambles under Expected Utility," *Economics Letters*, 2006, 91(2), 250-259.
- Pratt, John W., "Risk Aversion in the Small and in the Large," *Econometrica*, 32, 1964, 123-136.
- Rabin, Matthew, "Risk Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, 68, 2000, 1281-1292.

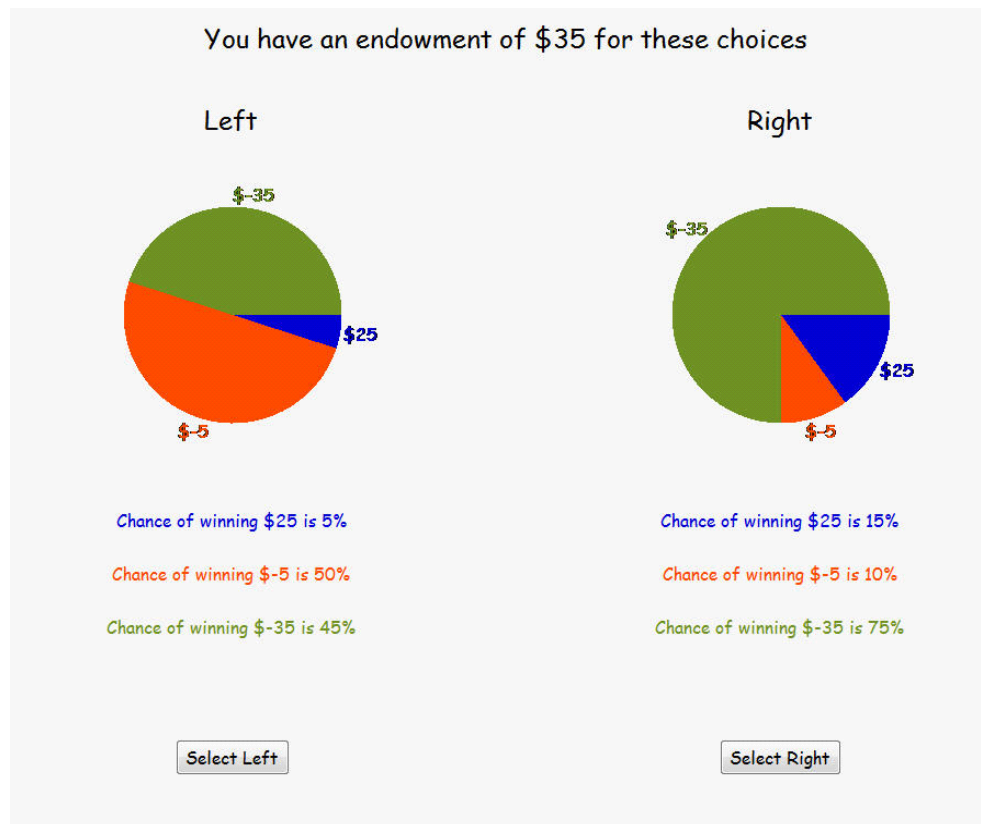
- Rabin, Matthew, and Thaler, Richard H., “Defending Expected Utility Theory: Response,” *Journal of Economic Perspectives*, 2002, 16(2), 229–230.
- Safra, Zvi, and Segal, Uzi, “Calibration Results for Non-Expected Utility Theories,” *Econometrica*, 76(5), 2008, 1143-1166.
- Schechter, Laura, “Risk aversion and expected-utility theory: A calibration exercise,” *Journal of Risk & Uncertainty*, 35, 2007, 67-76.
- Vickrey, William, “Measuring Marginal Utility by Reactions to Risk,” *Econometrica*, 13(4), October 1945, 319-333.
- von Neumann, John, and Morgenstern, Oskar, *Theory of Games and Economic Behavior* (Princeton: Princeton University Press, Third Edition, 1953).
- Watt, Richard, “Defending Expected Utility Theory: Comment,” *Journal of Economic Perspectives*, 2002, 16(2), 227–229.
- Wilcox, Nathaniel T., “Is the Premise of Risk Calibration Theorems Plausible?” *Presentation*, CEAR Workshop, Durham University, September 17, 2013.

Appendix A: Instructions (NOT FOR PUBLICATION)

Choices Over Risky Prospects

This is a task where you will choose between prospects with varying prizes and chances of winning each prize. You will be presented with a series of pairs of prospects where you will choose one of them. There are 100 pairs in the series. For each pair of prospects, you should choose the prospect you prefer. You will actually get the chance to play **one** of these prospects, and you will be paid according to the outcome of that prospect, so you should think carefully about which prospect you prefer.

Here is an example of what the computer display of such a pair of prospects will look like.



The outcome of the prospects will be determined by the draw of a random number between 1 and 100. Each number between, and including, 1 and 100 is equally likely to occur. In fact, you will be able to draw the number yourself using two 10-sided dice.

You will be told your cash endowment for each lottery at the top of the lottery. In this example it is \$35, so any earnings would be added to or subtracted from this endowment. The endowment may change from choice to choice, so be sure to pay attention to it. The endowment you are shown only applies for that choice.

In this example the left prospect pays twenty-five dollars (\$25) if the number drawn is between 1 and 5, and pays negative five dollars (\$-5) if the number is between 6 and 55, and pays negative thirty-five dollars (\$-35) if the number is between 56 and 100. The blue color in the pie chart corresponds to 5% of the area and illustrates the chances that the number drawn will be between 1 and 5 and your prize will be \$25. The orange area in the pie chart corresponds to 50% of the area and illustrates the chances that the number drawn will be between 6 and 55 and your prize will be \$-5. The green area in the pie chart corresponds to 45% of the area and

illustrates the chances that the number drawn will be between 56 and 100. When you select the lottery to be played out the computer will confirm what die rolls correspond to the different prizes.

Now look at the pie on the right. It pays twenty-five dollars (\$25) if the number drawn is between 1 and 15, negative five dollars (\$-5) if the number is between 16 and 25, and negative thirty-five dollars (\$-35) if the number is between 26 and 100. As with the prospect on the left, the pie slices represent the fraction of the possible numbers which yield each payoff. For example, the size of the \$25 pie slice is 15% of the total pie.

Even though the screen says that you might win a negative amount, this is actually a loss to be deducted from your endowment. So if you win \$-5, your earnings would be $\$30 = \$35 - \$5$.

Each pair of prospects is shown on a separate screen on the computer. On each screen, you should indicate which prospect you prefer to play by clicking on one of the buttons beneath the prospects.

After you have worked through all of the pairs of prospects, raise your hand and an experimenter will come over. You will then roll two 10-sided dice to determine which pair of the 100 prospects you chose will be played out. Since there is a chance that any of your 100 choices could be played out for real, you should approach each pair of prospects as if it is the one that you will play out. Finally, you will roll the two ten-sided dice again to determine the outcome of the prospect you chose.

For instance, suppose you picked the prospect on the right in this example, and it was the pair chosen to be played. If the random number from your rolls of the dice was 7, you would win \$25 in addition to your endowment; if it was 93, you would lose \$35 from your endowment. If you picked the prospect on the left and drew the number 7, you would lose \$5 from your endowment; if it was 93, you would again lose \$35 from your endowment.

Therefore, your payoff is determined by three things:

- which prospect you selected, the left or the right, for each of these 100 pairs;
- which prospect pair is chosen to be played out in the series of 100 such pairs using the two 10-sided dice; and
- the outcome of that prospect when you roll the two 10-sided dice again.

Which prospects you prefer is a matter of personal choice. The people next to you may be presented with different prospects, and may have different preferences, so their responses should not matter to you or influence your decisions. Please work silently, and make your choices by thinking carefully about each prospect.

All payoffs are in cash, and are in addition to the \$5 show-up fee that you receive just for being here, as well as any other earnings in other tasks from the session today.